analytical chemistry

Article

Molecular approaches to address the challenges of RNA analysis in complex matrices

Shan S. Lansing, Susmitha Matlapudi, and Sean M. Burrows

Anal. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.analchem.8b01621 • Publication Date (Web): 03 Jul 2018

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

For submission to Analytical Chemistry

Molecular approaches to address the challenges of RNA analysis in complex matrices

By

Shan S. Lansing, Susmitha Matlapudi, and Sean M. Burrows*

*corresponding author: sean.burrows@oregonstat.edu

Department of Chemistry

153 Gilbert Hall

Corvallis, OR 97330

Abstract:

We present on a design change and addition of an internal polyethylene glycol (PEG) spacer to an existing biosensor. There were two reasons for changing the sensor design. The first was to increase the stability of the biosensor against binding off-analytes with single nucleotide polymorphisms. The second was to prevent sensor degradation by nucleases. The biosensor, designed for detection of short non-coding RNA strands, is composed of Reporter and Probe nucleic acid strands that form a partially-complementary duplex. The internal PEG was added to the Reporter, and subsequently diminished false negatives that resulted from off-oligonucleotides binding. Furthermore, the PEG eliminated degradation of the sensor by DNase1 endonuclease. Currently, in situ and crude cell lysate RNA analysis is hindered by nonspecific interactions and degradation by endogenous nucleases. Together, the design changes presented here mitigate these matrix effects and allow for robust RNA analysis in complex medias.

Introduction:

Current methods to analyze RNA or protein rely on accurately and efficiently detecting small changes in expression[1,2]. Measuring changes in expression of RNA is interesting because RNA levels change prior to the expression of proteins[3,4]. The need to monitor RNA changes *in situ* exists because expression depends on the cell state and type. Furthermore, healthy and diseased tissue exist as heterogeneous collections of cells with different RNA expressions throughout the cell population[4,6]. Monitoring RNA with quantitative-real-time-Polymerase Chain Reaction (qRT-PCR) and microarrays requires extracting and purifying all RNA from cells[5,7]. However, location specific information is lost and low RNA expression can be diluted below the detection limit[8]. Thus, developing nucleic acid (NA) based sensors with the potential for intracellular analysis is an active field of research[9].

Programmable Watson-Crick base pairing, low cost of synthesis, ease of conjugation with reporter molecules, and intrinsically high affinity for complementary nucleotides make NA sensors attractive[10–12]. Promising *in vitro* results indicating the power of NA sensors, include: 1) femto- to nanomolar limits of detection, 2) wide linear range, and 3) short time-to-result[9,13]. However, applying NA sensors *in situ* or in crude cell lysate is limited by false signals from off-analyte oligonucleotides that interact with the sensor and by endogenous nucleases that degrade the sensor[14–17].

Small non-coding RNA range in size from approximately 20 to 100 nucleotides, and examples include: microRNA, small interfering RNA, piwi-RNA, and small nuclear RNAs. Expression of these transcripts is dynamic, and copy numbers can range from 1 to 50,000 copies[9]. This corresponds to a concentration range of approximately 2 pM to 80 nM in a cell of 1 pL volume. Irregular expression of small ncRNAs has been recorded in several cases (e.g. healthy cells, cancers, neurological disorders, and cardiovascular disease)[18-28], making them a relevant biomolecule to study. Current NA sensing

technologies are best for long nucleotide strands (predominantly messenger RNA)[1]. Due to short length of small RNA, current sensor designs often struggle to differentiate between one or two base pair mismatches and reach pico- to nanomolar sensitivity with minimal false signals.

Sensing small RNA strands is challenging due to the size and sequence similarity of RNAs in the same family[9]. Some small ncRNA (e.g. , miRNA) have family members that often express simultaneously and only differ by a few bases, called single nucleotide polymorphisms (SNPs)[8]. NA sensors generally rely on complementary hybridization between a portion of the sensor and the nucleotide analyte; and depends on thermodynamics driving the system toward analyte binding. One thermodynamic challenge, specific to sensing short RNA strands, is that complexes formed between the sensor and analyte have a wide range of melting temperatures, dependent on the G/C content of the analyte[9]. This, coupled with the presence of SNPs of the analyte, requires innovative technologies to create more robust NA sensors for short nucleotide analytes in complex biological medias[7].

The current gold standards for *in situ* analysis of nucleotide strands are Fluorescent *In Situ* Hybridization (FISH)[29], molecular beacons[30], and SmartFlares[31]. the Supporting Information (SI Introduction) has a description of these technologies and places for improvement. Molecular beacons and SmartFlares™ rely on disrupting a quencher-dye pair to turn the signal ON. This is problematic because there are many nonspecific matrix effects that can separate the dye and quencher, leading to false signals. To address these issues, we have previously developed a strand-displacement biosensor that forces Förster Resonance Energy Transfer (FRET) dye-pairs together to turn the signal ON[32]. The FRET-enhancement transduction mechanism is advantageous over the quenching mechanism because there are fewer matrix effects that will force a donor- and acceptor-dye together.

Although our biosensor has made advancements in the field of small ncRNA sensing[9], additional challenges must be overcome before it is ready for *in situ* application or work in raw-cell lysate. Specifically, changes must be made to mitigate off-oligonucleotide interactions and degradation by endogenous nucleases[33]. The challenges of measuring nucleic acids *in situ* and within complex cellular matrices are well documented and must be addressed[11,33].

In this work, we systematically investigated the performance of two new sensors in terms of off-analyte binding, enzyme degradation, and limit of detection. A previous Reporter+Probe sensor design served as a control to determine the efficacy of the design and chemical changes. We changed the design by adding a stem to the Probe strand, on the 3' end. The chemical change made was to replace non-functional DNA with an internal polyethylene glycol (PEG) spacer on the Reporter strand. Given the biosensor has two oligonucleotide strands, we considered off-oligo interactions for both the Probe and Reporter strands. The new biosensors were made for miR29b-1-5p, as model to test the sensor's design. We will discuss how changes to the chemical make-up and molecular architecture were instrumental in creating a more stable biosensor. To the best of our knowledge, this study is the first to incorporate an internal PEG spacer for the purpose of mitigating false signals due to nuclease degradation and nonspecific binding.

Experimental:

Oligonucleotide Sequences and Sample Preparation

All oligonucleotides were purchased from Integrated DNA Technologies, except the Probe. The Probe was purchased from Exiqon. Oligonucleotides were diluted in a phosphate buffered saline (PBS), containing 10 mM Tris buffer, 2.5 mM $MgCl_2$, and 0.005% Tween 20 (obtained from Fisher Scientific, used as received); for simplicity, this solution will now be

referred to as a 'PBS solution'. All solutions containing Reporter and Probe oligonucleotides were prepared to contain a final concentration of 100 nM. For a more detailed description of the sample preparation and concentration verification see SI Experimental.

Fluorimeter Set Up and Signal Processing

In-depth descriptions of the fluorimeter, spectroscopic apparatus, and data acquisition are provided in SI Experimental.

Limit of Detection Analysis in Cell Lysate

To determine if the internal PEG spacer affected the RNA-sensing capabilities of the biosensor, calibration curves were constructed for both the DNA-Sensor and the DNA+PEG-Sensor in MCF-7 cell lysate. Details on how the samples were prepared are in the SI Experimental.

Off-analyte interactions with Probe or Reporter strand

This biosensor is designed to detect a DNA copy of mmu-miR29b-1-5p. (DNA was used due to increased stability during storage, mmu = *Mus musculus*, mouse.) The single nucleotide polymorphism (SNP) off-analyte was mmu-miR29b-2-5p, a biologically relevant miRNA in the same family as the analyte (see Table S1A to compare the sequences).

Susceptibility to false positives was tested by comparing the Cy5 signal change in response to miR29b-2-5p or a mixture of an Off-Analyte Cocktail (containing miR26a, miR27a, miR146a, miR146b, and miR146a-3p). The off-analytes in the 'Off-Analyte Cocktail' have no sequence similarity to the analyte. A detailed description of the sample prepration can be found in SI Experimental.

The Reporter+Probe biosensor may be susceptible to off-oligonucleotides interacting with the Reporter. In order to determine the Reporter's susceptibility to off-oligo interactions, an oligonucleotide (referred to as "12-mer") that was partially complementary to the Reporter was purchased. Susceptibility to false negatives was tested by comparing the Cy5 signal change in response to the 12-mer binding to Reporter.

False Negatives from Nuclease Degradation

To analyze the sensor's ability to withstand nuclease degradation, we compared the fluorescent signal with and without nuclease added. Two endonucleases were tested: S1 endonuclease and DNase1. After endonuclease treatment, the fluorescent signal of each sample was then collected at room temperature.

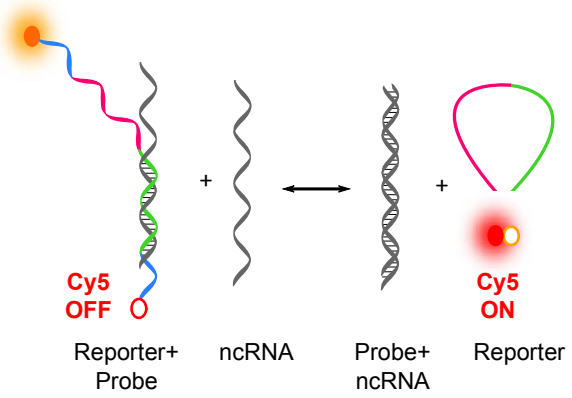
Specific details regarding sample preparation with DNase1 and S1 endonuclease is provided in SI Experimental. Briefly, samples of 100 nM Reporter, 100 nM Reporter+Probe, and 100 nM Reporter+Probe+miR29b-1-5p were examined both with and without each endonuclease added. The S1 endonuclease was purchased from Promega and stored at -20 °C. Based on work by Chen *et al.*[14], 0.05 U/μL of S1 endonuclease was added to each solution and final volume was brought to 200 μL in 1x S1 Buffer.

DNase1 endonuclease was purchased from New England BioLabs Inc. and stored at -20°C. Sample treatment was similar to that of S1, but the solutions were treated to 250 U/μL DNase1 endonuclease and the final volume was brought to 200 μL. The final concentration of DNase1 was based on work by Yang *et al.*[35].

Theory:

Changes to Probe Design and Reporter Chemistry

The biosensor here is composed of two partially complementary DNA strands, referred to as the Reporter and the Probe. The Probe is fully complementary to the analyte strand. The biosensor's recognition mechanism relies on the analyte binding to a region on the Probe (called the toehold region) and displacing the Reporter from the Probe (see Scheme 1). On distal ends of the Reporter are fluorescent dyes capable of Förster Resonance Energy Transfer (FRET). In the Reporter+Probe complex, the dyes on the Reporter are far apart forcing the acceptor-dye (Cy5) into the OFF state, emitting a very weak signal. The Reporter is displaced from the probe by the analyte and folds into a hairpin. When the Reporter forms a hairpin, the dyes are brought within the FRET distance. In the hairpin conformation, the donor-dye transfers energy to the acceptor-dye (turning the acceptor-dye signal ON).



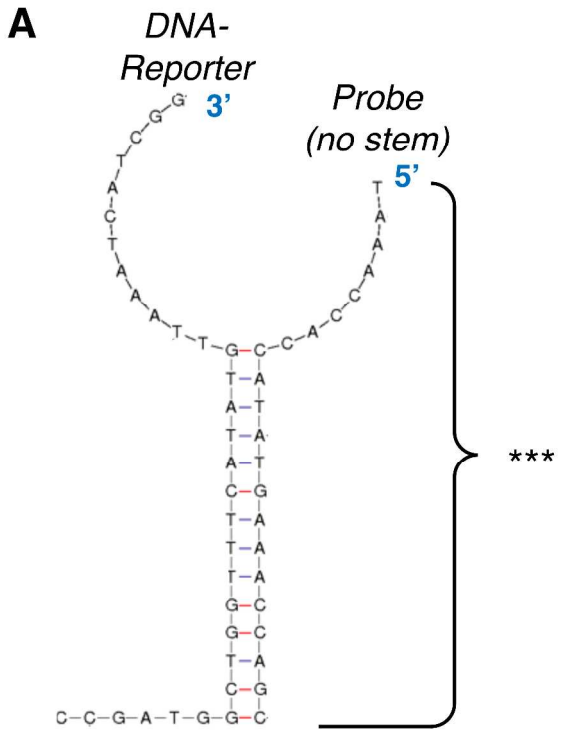
Scheme 1: Biosensor's Recognition and transduction mechanism. The biosensor has a Reporter and a Probe DNA-strand. Cy3 is excited by 945 nm light that does not excite Cy5 well. In the Cy5 OFF state, the Cy3 dye is excited but is too far from Cy5 to transfer energy. The analyte binds to the toehold region in the Probe (unbound gray portion). The analyte (gray) displaces the Reporter from the Probe. The Reporter then folds into a hairpin, bringing the distal dyes together and within the FRET distance, Cy3 transfers energy to Cy5, and the Cy5 signal turns ON. Reporter colors: Pink = non-complementary region, Green = complementary to probe, blue = hairpin-stem complementary regions.

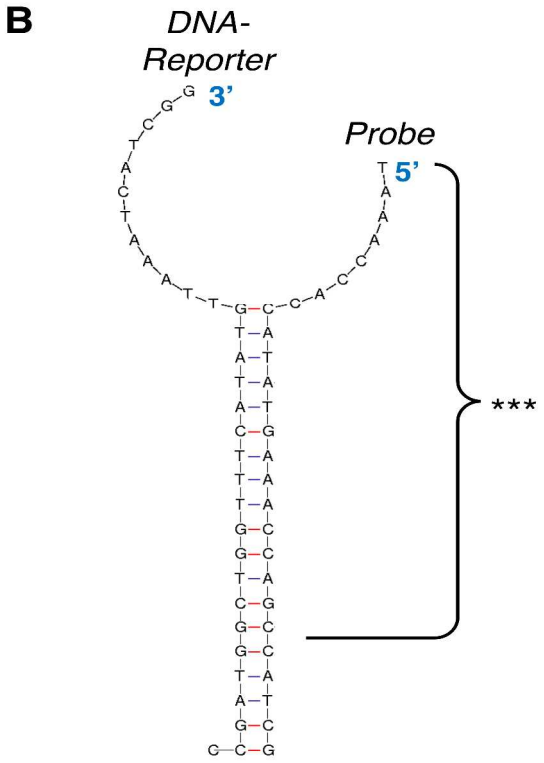
The Probe+analyte complex must be more stable than the Reporter+Probe complex, so that Scheme 1 will be driven towards the products. The formation of Reporter+Probe is dependent on both thermodynamics and kinetics, though the focus of this work will be primarily on thermodynamic stability and experimental data driven by the thermodynamic analysis. The stability of the Probe+analyte complex is determined by the sequence of the analyte. In order to increase the ΔG between Reporter+Probe to Probe+analyte, four Locked Nucleic Acids (LNAs) were strategically added to the Probe in locations that only bind to the analyte, and not the Reporter. LNA base pairs are known to have a higher binding affinity than canonical NAs, meaning a more stable Probe+analyte complex[36]. By selectively placing the LNAs in regions of the Probe that binds the analyte, we were able to increase the favorability of Probe+analyte formation over Reporter+Probe. The sequences and predicted thermodynamics of the Probe, Reporter, and other oligonucleotides for this study are found in Table S1B of the Supporting Information.

The first change made to the biosensor was adding a DNA segment on the 3' end of the Probe that was complementary to all but one nucleotide in the Reporter stem (but not the analyte). the stability of the Reporter+Probe complex was increased by adding this DNA segment. This region, termed 'Probe stem', allowed for greater control over the thermodynamics of the Reporter+Probe complex.

The Probe was previously designed to be an exact reverse complement of the analyte[37], referred to here as $Probe_{No\ stem}$. Figure 1A and Scheme 1 show Reporter+$Probe_{No\ stem}$. The design change made to this sensor can be seen in Figure 1B ('DNA-sensor' is the Reporter+Probe, or 'DNA-Reporter' when it is the Reporter strand alone). We named the Reporter 'DNA-Reporter' to differentiate it from the chemically modified 'DNA+PEG Reporter' described below.

Ultimately, $Probe_{No\ stem}$ was never synthesized because the predicted equilibrium concentration was only 44.13 nM at 37°C for a mixture of 100 nM Reporter with 100 nM $Probe_{No\ stem}$ (Figure 2A). By adding a Probe stem the predicted concentration of Reporter+Probe increased to 99.66 nM (Figure 2B).


A
DNA-
Reporter
3'
Probe
(no stem)
5'
***


B
DNA-
Reporter
3'
Probe
5'
***

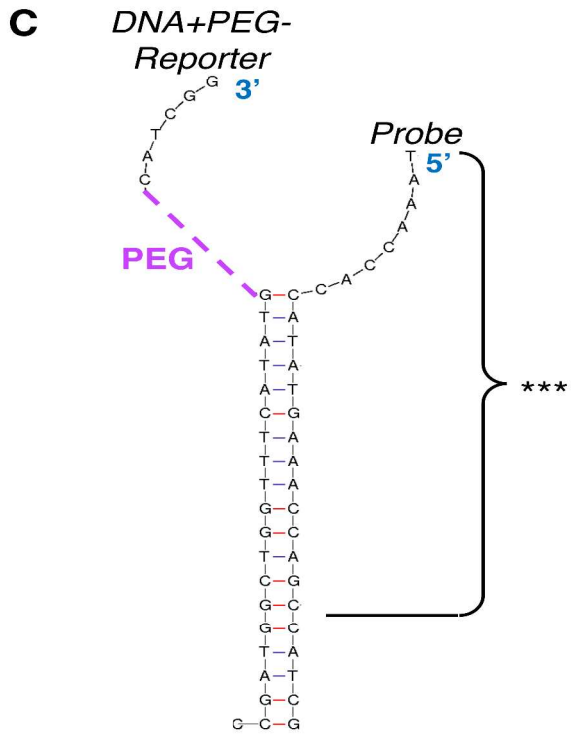
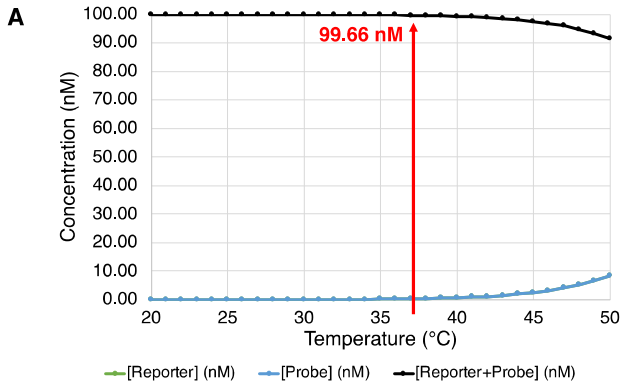


Figure 1: Biosensor designs. The *** denotes the region of the Probe that is complementary to the analyte. **(A)** The biosensor design from previous work. The Probe was partially complementary to the Reporter and entirely complementary to the analyte. **(B)** The first change was adding a 3' Probe-stem that was complementary to the Reporter's stem. The Probe stem was not complementary to the analyte. **(C)** The second change was to replace the non-complementary region on the Reporter with a PEG spacer (purple dashes).



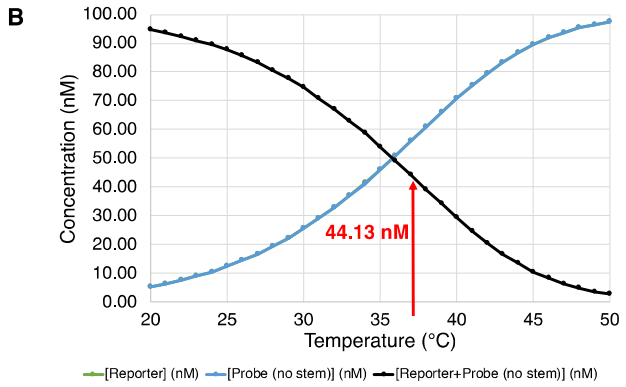


Figure 2: Predicted equilibrium concentrations of solutions containing 100 nM Reporter and 100 nM Probe, (A) without and (B) with a Probe stem. The Probe stem, brought the biosensor's predicted equilibrium concertation from 44.13 % Reporter+ $Probe_{No\ stem}$, to 99.66% Reporter+$Probe_{Stem}$. The predicted concentration of the biosensor ([Reporter+Probe]) is the black line. The total predicted concertation of unbound Reporter and Probe are indicated by green and blue lines, respectively. Unbound [Reporter] is approximately equal to unbound [Probe], so the green line for unbound [Reporter] cannot be seen on the graphs.

Instead of adding a Probe stem, the stability of the Reporter+Probe complex could have been improved by increasing the number of complementary base pairs between the Reporter strand and the Probe strand. However, when the number of complementary bases between the Reporter and Probe are increased, the toehold region on the Probe is shortened. When the toehold region was shorted, the Probe had three fewer open bases to initiate binding to the unbound analyte. Therefore, decreasing the number of bases in the toehold region may reduce the analyte's ability to displace the Reporter.

An example of a Reporter+Probe complex that is stable at 37°C without a Probe stem can be found in Figure S1. Although the predicted concentration of this Reporter+Probe increased to 99.60 nM, the toehold region on the Probe was reduced to 4 bases, all of which were weak (A/T). Shortening the toehold region hinders the recognition mechanism because the analyte may not bind to the Probe. Conversely, by adding a Probe stem we were able to increase the predicted equilibrium concentration and maintain the length of the toehold region.

The second change made to the biosensor was to replace the Reporter's non-complementary region with an internal polyethylene glycol (PEG) spacer. The region of the biosensor known as the 'non-complementary region' is not part of the hairpin's stem and is not complementary to the Probe. In Scheme 1, the non-complementary region is shown in pink. The sequence of the non-complementary region is determined by an in-house MATLAB code, described in previous work[38]. For details on the code and predicted thermodynamics from the RNA Institute at SUNY Albany see the SI Theory section[39].

We hypothesized that an internal PEG spacer in the non-complementary region would reduce off-oligo interactions and disguise it from enzymes. The location of the internal PEG spacer for this sensor (referred to as 'DNA+PEG-Sensor' or 'DNA+PEG-Reporter') can be seen in Figure 1C as a dashed purple line.

In this study, we compared two biosensors for mmu-miR29b-1-5p that each have a stem on the Probe strand that is complementary to the Reporter stem (Figure 1B, C). One Reporter strand was composed entirely of DNA, and the other Reporter strand had an internal PEG spacer in between DNA segments. The region of the Reporter that was replaced with a PEG spacer had no intended binding-function, thus the addition of a PEG was not expected to impede RNA sensing ability.

Predicted Probe Selectivity

Off-analyte (OA) interactions occur when the Probe+OA complex is more stable than the Probe+Reporter, such that the OA can displace the Reporter from the Probe. By adding a Probe stem there is more control over the thermodynamics of the Reporter+Probe complex, ultimately allowing for increased predicted stability and selectivity (more information is in SI Theory).

As shown in Table 1, $Probe_{No\ stem}$+miR29b-2-5p was slightly more stable than Reporter+$Probe_{No\ stem}$ (miR29b-2-5p is a SNP from the same family as the intended analyte). This indicates that miR29b-2-5p may displace the Reporter from the $Probe_{No\ stem}$ and result in a false positive. However, by adding a stem on the Probe, the Reporter+$Probe_{stem}$ was significantly more stable than the $Probe_{stem}$+miR29b-2-5p complex. We chose to continue with the added stem on the Probe because this biosensor is expected to discriminate against single nucleotide polymorphism (SPN) off-analytes.

Table 1. Predicted binding thermodynamics of the Probe strand, with and without a Probe stem

	ΔG (kcal/mol)	**T_m (°C)**
'Probe' has NO stem		
Reporter+$Probe_{No\ stem}$	-14.4	46.8
(OA): $Probe_{No\ stem}$ +miR29b-2-5p	-15.2	49.7
(A): $Probe_{No\ stem}$ +miR29b-1-5p	-23.3	62.0
'Probe' has stem at the 3' end		
Reporter+Probe	-21.4	58.9
(OA): Probe +miR29b-2-5p	-15.2	49.7
(A): Probe+miR29b-1-5p	-23.5	61.2

*OA = off-analyte; A = Analyte.

Results and Discussion:

Probe Selectivity

Scheme 1 shows that the Cy5 signal is OFF in the Reporter+Probe conformation and turns ON when analyte is added. A false signal would have resulted if the Cy5 signal turned ON in the presence of an oligonucleotide strand that was not the analyte.

The analyte, mmu-miR29b-1-5p, has 1 isoform, mmu-miR29b-2-5p, that is missing a 5-prime Guanine (G) and has two internal single nucleotide polymorphisms (SNPs) (see SI Table S1A for sequence comparison). As expected from *in silico* predictions, the experimental results confirmed the Reporter+Probe had selectivity against SNPs (Figure S3). Figure S3A shows that the DNA-Sensor had a large change in Cy5 signal in the presence of the analyte (mmu-miR29b-1-5p), and no significant change in signal in the presence of the SNP off-analyte (mmu-miR29b-2-5p). Similar results were obtained for the DNA+PEG-Sensor (Figure S3B). In addition, neither sensor responded to cocktails of non-related sequences.

The data shows that both the DNA-Sensor and the DNA+PEG-Sensor were not susceptible to false positives from off-analytes, even from SNP sequences. These results are promising because it is necessary to develop highly selective biosensors for *in situ*, raw-cell lysate, and crude tissue analysis because multiple similar RNA strands can be expressed at the same time.

Reporter Selectivity

Unlike the Probe, predicting specific off-oligonucleotides that would interact with the Reporter was more difficult because it was not designed to bind to a transcript that would have known isoforms. For this reason, we created an off-oligonucleotide strand (termed "12-mer" due to its length) that partially bound to the Reporter. The area that an endogenous oligonucleotide would likely interact with the Reporter to cause false signals is the region with exposed nucleotides. Thus, the sequence of the 12-mer was the complement to the Reporter's non-complementary region and the Reporter's 3-prime stem (pink+blue Reporter regions in Scheme 1).

Figure 3 shows that the Cy5 signal decreased ($p < 0.05$) when the 12-mer was added to the DNA-Reporter. The decrease in signal symbolizes a false negative (i.e. first the Reporter hairpin forms due to the presence of the analyte, but then the Reporter was re-opened by the 12-mer and the signal turns OFF as if the analyte was not present. Conversely, when the 12-mer was added to the DNA+PEG-Reporter the Cy5 signal did not change ($p < 0.05$, Figure 3). The results indicate that the internal PEG spacer mitigated false negatives that result from off-oligomers binding to the Reporter strand.

Although the 12-mer was not a known short RNA transcript, the specific sequence of the 12-mer occurs 3 times in the mouse transcriptome (reference transcriptome provided by Ensembl)[42]. When considering a sequence of only 10 of the 12 nucleotides (i.e. the first 10, the middle 10, or the last 10 of the 12-mer), the number increased to a total of 218 occurrences. Although this data does not give information regarding the frequency or expression of those transcripts, it did provide evidence that the 12-mer was a valid concern for off-oligonucleotide interactions with the Reporter. Additionally, it is widely accepted (primarily for PCR primers) that approximately 10 base pairs are necessary for two strands to anneal. Therefore, the 12-mer strand represents a real concern for off-oligo interactions in biological matrices.

To ensure that the Probe could not be displaced from the Reporter by the 12-mer, we added the 12-mer to the respective Reporter+Probe samples. When the 12-mer was added to DNA+PEG-Reporter+Probe sample there was not a change in signal. Conversely, the DNA-Reporter+Probe sample resulted in a decrease in Cy5 signal (Figure S2). This result was surprising because Reporter+Probe conformation was the OFF state (or baseline) level of the biosensor.

The decrease in Cy5 signal of the DNA-Reporter+Probe after the added 12-mer was most likely due to residual Reporter strands existing in the hairpin conformation instead of the Reporter+Probe complex. Although over 99% Reporter+Probe was predicted to form (Figure 2B), these experimental results indicated that there was some fraction of Reporter strands still in the hairpin conformation (ON) that could be opened (OFF) to decrease the signal further. By incorporating a PEG spacer into the Reporter strand we were able to avoid such false signals from off-oligos binding to the Reporter hairpin.

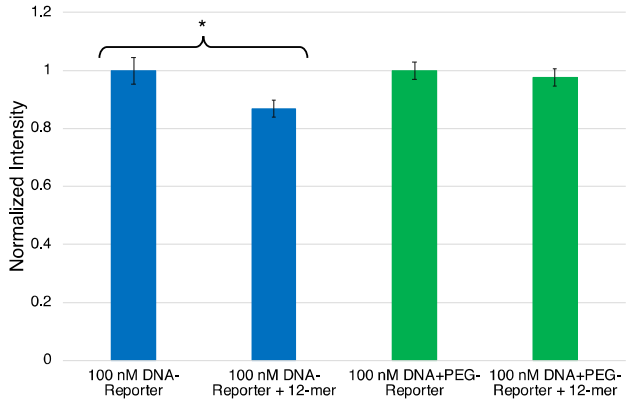


Figure 3: Interference by off-oligonucleotide '12-mer' binding to Reporter strand. The Cy5 signal of the DNA-Reporter decreased due to the presence of the 12-mer (blue bars, * $p<0.05$, N = 3). Conversely, the Cy5 signal of the DNA+PEG-Reporter did not change in the presence of the 12-mer (green bars). Adding the PEG spacer resulted in a more robust biosensor that reduces susceptibility to off-oligomer related false signals. The signals were normalized to the Cy5 signal of 100 nM Reporter for both the DNA-Reporter system and the DNA+PEG-Reporter system, respectively. Error bars represent the standard deviation between trials. Samples were excited at 945 nm.

Nuclease Degradation

For future *in situ* work to be successful, nucleic acid biosensors must evade damage from nucleases. We compared the newly designed DNA-Sensor and DNA+PEG-Sensor, to our previous biosensor design (Sensor-25)[41]. Sensor-25 was compared to the new sensor designs

because Sesnor-25 did not have the Probe stem nor an internal PEG spacer on the Reporter. Table 2 outlines the differences among the three biosensors.

Sensor-25 (previously named 'miR-26a-R1') was designed for a different analyte than the DNA-Sensor and DNA+PEG-Sensor[38]. For this reason, it was not included in the selectivity studies. However, because the analyte does not majorly affect how nucleases degrade sensors, Sensor-25 served as a reference point to an unmodified design. All three sensors have Locked Nucleic Acids (see SI Table S1B). Of note, the change in intensity from OFF state to ON state is noticeably smaller for Sensor-25 than for either DNA-Senor or DNA+PEG-Sensor (ΔI based on data without added nuclease from Figure 4).

Table 2. Notable differences between the three biosensor designs.

	Sensor-25	**DNA-Sensor**	**DNA+PEG-Sensor**
Probe stem	No	Yes	Yes
Chemical Make-up	DNA	DNA	DNA & PEG
Temperature for sensing miRNA	25 °C	37 °C	37 °C
LNAs present	Yes	Yes	Yes
Analyte	mmu-miR26a-*2*-3p	mmu-miR29b-1-5p	mmu-miR29b-1-5p
$\Delta I_{OFF \rightarrow ON}$ (counts)	7605 ± 639	42157 ± 1721	41352 ± 7538

First, we will discuss sensor how the sensors hold up to S1 endonuclease , which typically targets single stranded DNA (ssDNA)[43] – such as the Reporter hairpin. If the Reporter hairpin was degraded by S1 endonuclease, then we expect the Förster Resonance Energy

Transfer (FRET) between the dyes to be disrupted, and the Cy5 signal would decrease because it is no longer receiving energy from Cy3.

To determine if there was a difference between the extent to which S1 endonuclease degraded the 3 biosensors, the changes in signal due to nuclease degradation were compared – the results are shown in Figure S4. For all sensors, there was no significant change of signal for the Reporter hairpin or the Reporter+Probe samples. However, all 3 sensors showed a significant decrease in Cy5 signal for the Reporter+Probe+Analyte sample (Figure S4).

The decrease in Cy5 signal indicates that the donor and acceptor dye were separated due to Reporter hairpin being degraded by S1 nuclease. These results were surprising because the Reporter strand exists in a hairpin configuration for both the 100 nM Reporter and 100 nM Reporter+Probe+Analyte samples. Because both samples have reporter hairpins, we expected that both (or neither) samples would show evidence that the Reporter hairpin was degraded. However, the results indicate that the Reporter hairpin was degraded in the 100 nM Reporter+Probe+Analyte sample but not in the 100 nM Reporter sample. These unexpected results and a possible explanation are detailed in Figure S4 and caption.

Although unexpected, the observed nuclease related damage of the Reporter+Probe+Analyte samples was consistent for all biosensor designs. Among the 3 biosensors, we did not observe any statistically significant difference in the average percent change of signal intensity (%ΔI) due to S1 degrading the sensors. For Reporter-25, DNA-Reporter, and DNA+PEG-Reporter, the average %ΔI of the Reporter+Probe+Analyte samples was -14.4 ± 2.97%, -13.5 ± 3.56%, and -18.4 ± 1.49%, respectively (Figure S5, a detailed description of the (%ΔI) calculation is in SI Results and Discussion). These results indicate that

changing the design and chemical make-up of the sensor did not mitigate or exacerbate degradation by the endonuclease.

Next, we determined the extent to which DNase1 endonuclease would degrade the 3 biosensor designs. DNase1 has ~ 500 times greater affinity for double stranded DNA (dsDNA) over ssDNA[31]. The Reporter+Probe complex is primarily dsDNA, and we have previously reported on false positive and negative signals due to DNase1 damage to the Reporter+Probe[37].

In the previous Reporter+Probe$_{\text{No stem}}$ design (Figure 1A), the Reporter hairpin stems extend beyond the Probe as ssDNA; we hypothesized that when the Reporter+Probe complex was degraded by a dsDNA nuclease, the Reporter stems survived and were able to rehybridize and change the signal[37]. By adding a 3' Probe stem (complementary to the 5' Reporter stem) we expected to reduce the possibility of DNase1 degrading the biosensor, but sparing the stems of the Reporter hairpin.

As expected, Sensor-25 showed a significant change in signal for all Reporter conformations (Figure 4A), indicating that DNase1 degraded the Reporter hairpin and the Reporter+Probe of Sensor-25. The Cy5 signal for Reporter-25+Probe (OFF state) decreased after DNase1 was added. In contrast to previous data, the stems of Sensors-25 did not appear to re-hybridize because if they did, then the Cy5 signal would have increased after the dsDNA portion of Reporter-25+Probe was degraded.

For the other solutions containing the Reporter hairpin of Sensor-25 (100 nM Reporter-25 and 100nM Reporter25+Probe+Analyte solutions) the decrease in Cy5 signal was expected because the DNase1 can degrade the dsDNA hairpin stem and separate the donor and acceptor dye. Ultimately these changes in signal represented a false negative because the Reporter hairpin

would form due to the presence of the analyte, and the signal would be turned back OFF after DNase1 degradation of the Reporter hairpin.

There was no significant change in Cy5 signal due to DNase1 degrading the DNA-Reporter+Probe complex (Figure 4B). In solutions containing the Reporter hairpin (100 nM Reporter-25 and 100nM Reporter25+Probe+Analyte solutions), the decrease in signal indicates DNase1 degraded the DNA-Reporter hairpin. The incorporated PEG spacer further eliminated false signals from all test cases with DNase1 (Figure 4C). The data provides evidence that adding the Probe stem and internal PEG to the Reporter resulted in a more robust biosensor that was less susceptible to false signals from DNase1 damage.

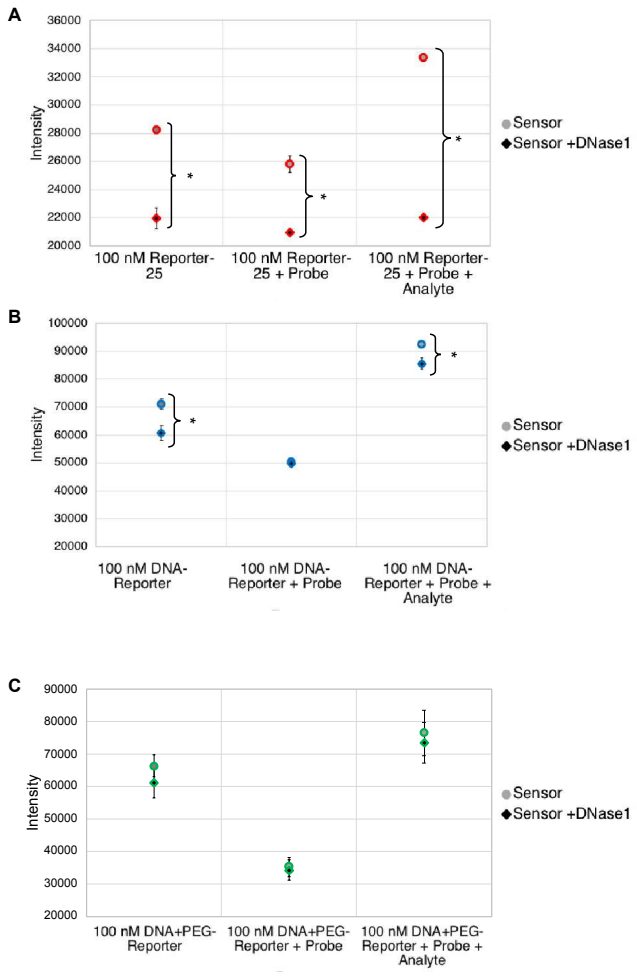


Figure 4: Signal response of 3 biosensor designs due to DNase1, mostly a dsDNA endonuclease. (* $p<0.05$, N = 3) Gray filled circles denote the signal from the sensor, and the black filled diamonds denote the sensor's signal after exposure to DNase1. **(A)** All Sensor-25 samples showed a decrease in the Cy5 intensity due to DNase1. Sensor-25, outlined in red, did not have a 3' Probe stem nor Reporter internal PEG. **(B)** DNA-Reporter+Probe samples, outlined in blue, showed no change in signal due to DNase1. This indicated that the addition of a Probe stem mitigated false signals from DNase1 degradation. However, solutions containing the DNA-Reporter hairpin, showed a decrease in Cy5 signal. **(C)** There was no change in signal of the DNA+PEG-Sensor solutions, outlined in green. Adding a Probe stem and Reporter PEG spacer, further mitigated DNase1 related false signals. Error bars represent the standard deviation between 3 trials, and some are hard to see. All samples were excited at 945 nm.

Limits of Detection

Given that Cy5 intensity increased as analyte was added, the limit of detection was calculated as:

$$I_{R+P} + 3(\sigma) = m(LOD) + b$$

Where 'I_{R+P}' was the intensity of Cy5 from the 100nM Reporter+Probe solution, 'm' was the slope of the calibration curve, 'b' was the y-intercept, and 'σ' was the standard deviation of I_{R+P}. The limits of detection and calibration curves for the DNA-Sensor and DNA+PEG-Sensor are in SI Table S2 and Figure S6. Both the sensors demonstrated ~ 6 nM detection limits in MCF-7 cell lysate. These LOD were consistent with LODs of our previous biosensors and those for similar biosensors in the literature[9] indicating the Probe stem and internal PEG spacer did not impact the biosensors' abilities to detect RNA. As previously noted, depending on the cell, microRNA are believed to range from 1.66 pM to 83 nM *in situ*. Although the LOD presented here is at the upper end of the biologically relevant range, we have shown in previous work that lower LODs are achievable with lower sensor concentration and longer integration times.

Conclusions:

The biosensor with both a Probe stem and internal PEG spacer showed complete

mitigation of false signals due to DNase1 endonucleases and relevant off-oligonucleotide bindings. The results from the DNA+PEG-Reporter+Probe biosensor (also known as DNA+PEG-sensor) demonstrate the feasibility for future work in complex biological samples. The findings presented here are part of an ongoing effort to design selective and robust biosensor for intracellular analysis of small non-coding RNAs (small ncRNA). Highly selective RNA biosensors are necessary for small ncRNAs because they belong to families that often have similar sequences. Additionally, there are numerous reports that have established that *in situ* capabilities of NA biosensors are limited by vulnerability to hydrolysis by endogenous nucleases[14,16,29,35]. Further studies are needed to determine the relevant nuclease profile of the intended cell lines for future analysis. This will be important for determining if the unmitigated degradation by S1 nuclease is a significant concern for potential work in biological samples.

Currently we are investigating methods to transfect the sensor into cells for future *in situ* sensing. Due to the hydrophobic structure of PEG, there is potential that the internal PEG spacer will enable the biosensor to diffuse more easily across the cell membrane. However, changing the biosensors hydrophobicity could possibility lead to unwanted interactions with other hydrophobic molecules *in situ*; this would be especially concerning in cells with a high lipid content, where the lipid molecules may coagulate around the sensor and impede interaction with the analyte. Further studied are needed to determine if addition of the PEG results in unwanted hydrophobic interactions.

We recently published on a multi-input nucleic acid (NA) biosensor that can detect a combination of 3 small ncRNAs[44]. The design principals and chemical changes made here can be applied to the multi-input sensor, in addition to other types of NA-based biosensors. In general, adding an internal PEG spacer to a biosensor constructed of NAs is likely to reduce false signals

from off-analyte interference and endonuclease damage making the sensors more robust and well suited for *in situ* analysis.

The advancements reported here are steps toward monitoring small ncRNAs in complex biological medias. Combined with the commercial availability and low cost of sequence-specific oligonucleotides, these results show promise to improve the applicability of NA-biosensors in crude cell lysate, crude tissue grounds, serum, and inside cells and tissue. The ability to sensitivity and specifically measure RNA expression will make it possible to better understand, diagnose, and track the progression of a myriad of diseases. As the demand for robust sensors continues to rise, it is highly probable that our contribution will be one of many developments that lead toward improved *in situ* sensing.

Supporting Information

Supporting Information available as indicated in text.

Acknowledgments:

The authors would like to thank the Research Office at Oregon State University for supporting this research through a grant from the General Research Fund. We would also like to acknowledge David Richardson for automating the process of biosensor sequence generation and *in silico* testing, in addition to design of the graphical abstract.

(1) Smith, S. J.; Nemr, C. R.; Kelley, S. O. *J. Am. Chem. Soc.* **2017**, *139* (3), 1020–1028.

(2) Tsien, R. Y. *Nat. Rev. Mol. cell Biol.* **2003**, *Suppl* (September), SS16-21.

(3) Croce, C. M. *Nat. Rev. Genet.* **2009**, *10* (10), 704–714.

(4) Andorfer, C. A.; Necela, B. M.; Thompson, E. A.; Perez, E. A. *Trends Mol. Med.* **2011**, *17* (6), 313–319.

(5) Parikh, Victoria N.; Chan, S. Y. *Methods Mol. Biol.* **2013**, *1024*, 157–172.

(6) Meacham, C. E.; Morrison, S. J. *Nature* **2013**, *501* (7467), 328–337.

(7) Johnson, B. N.; Mutharasan, R. *Analyst* **2014**, *139* (7), 1576.

(8) Cissell, K. A.; Shrestha, S.; Deo, S. K. *Anal. Chem.* **2007**, *79* (13), 4754–4761.

(9) Graybill, R. M.; Bailey, R. C. *Anal. Chem.* **2016**, *88* (1), 431–450.

(10) Chen, Y. J.; Groves, B.; Muscat, R. A.; Seelig, G. *Nat Nanotechnol* **2015**, *10* (9), 748–760.

(11) Juskowiak, B. *Anal. Bioanal. Chem.* **2011**, *399* (9), 3157–3176.

(12) Chakraborty, K.; Veetil, A. T.; Jaffrey, S. R.; Krishnan, Y. *Annu. Rev. Biochem.* **2016**, *85* (1), 349–373.

(13) Chugh, P.; Dittmer, D. P. *Wiley Interdiscip. Rev. RNA* **2013**, *3* (5), 17–36.

(14) Chen, A. K.; Behlke, M. A.; Tsourkas, A. *Nucleic Acids Res.* **2007**, *35* (16).

(15) Li, J.; Tan, S.; Kooger, R.; Zhang, C.; Zhang, Y. *Chem. Soc. Rev.* **2014**, *43* (2), 506–517.

(16) Nanoparticles, D. G.; Prigodich, A. E.; Alhasan, A. H.; Mirkin, C. A. *J. Am. Chem. Soc.* **2011**, No. DNase I, 2120–2123.

(17) Minunni, M. *Spectrosc. Int. J.* **2003**, *17* (2–3), 613–625.

(18) Taft, R. J.; Pang, K. C.; Mercer, T. R.; Dinger, M.; Mattick, J. S. *J. Pathol.* **2009**, *220* (October), 126–139.

(19) Esteller, M. *Nat. Rev. Genet.* **2011**, *12* (12), 861–874.

(20) Lu, J.; Getz, G.; Miska, E. A.; Alvarez-Saavedra, E.; Lamb, J.; Peck, D.; Sweet-Cordero, A.; Ebert, B. L.; Mak, R. H.; Ferrando, A. A.; Downing, J. R.; Jacks, T.; Horvitz, H. R.; Golub, T. R. *Nature* **2005**, *435* (7043), 834–838.

(21) Sassen, S.; Miska, E. A.; Caldas, C. *Virchows Arch.* **2008**, *452* (1), 1–10.

(22) Heegaard, N. H. H.; Schetter, A. J.; Welsh, J. A.; Yoneda, M.; Bowman, E. D.; Harris, C. C. *Int. J. Cancer* **2012**, *130* (6), 1378–1386.

(23) Alexandrov, P. N.; Dua, P.; Hill, J. M.; Bhattacharjee, S.; Zhao, Y.; Lukiw, W. J. *Int. J. Biochem. Mol. Biol.* **2012**, *3* (4), 365–373.

(24) Hébert, S. S.; De Strooper, B. *Trends Neurosci.* **2009**, *32* (4), 199–206.

(25) Yelamanchili, S. V; Chaudhuri, A. D.; Chen, L.-N.; Xiong, H.; Fox, H. S. *Cell Death Dis.* **2010**, *1* (9), e77.

(26) Kantharidis, P.; Wang, B.; Carew, R. M.; Lan, H. Y. *Diabetes* **2011**, *60* (7), 1832–1837.

(27) Xu, Q.; Seeger, F. H.; Castillo, J.; Iekushi, K.; Boon, R. A.; Farcas, R.; Manavski, Y.; Li, Y. G.; Assmus, B.; Zeiher, A. M.; Dimmeler, S. *J. Am. Coll. Cardiol.* **2012**, *59* (23), 2107–2117.

(28) Romaine, S. P. R.; Tomaszewski, M.; Condorelli, G.; Samani, N. J. *Heart* **2015**, *101* (12),

921–928.

(29) Song, R.; Ro, S.; Yan, W. *Methods Mol. Biol.* **2010**, *629*, 287–294.

(30) Tyagi, S.; Kramer, F. R. *Nat. Biotechnol.* **1996**, *14* (3), 303–308.

(31) Seferos, D. S.; Prigodich, A. E.; Giljohann, D. A.; Patel, P. C.; Mirkin, C. A. *Nano Lett.* **2009**, *9* (1), 308–311.

(32) Larkey, N. E.; Zhang, L.; Lansing, S. S.; Tran, V.; Seewaldt, V. L.; Burrows, S. M. *Analyst* **2016**, *141* (22).

(33) Kilic, T.; Erdem, A.; Ozsoz, M.; Carrara, S. *Biosens. Bioelectron.* **2018**, *99* (May 2017), 525–546.

(34) Kukol, A.; Li, P.; Estrela, P.; Ko-Ferrigno, P.; Migliorato, P. *Anal. Biochem.* **2008**, *374* (1), 143–153.

(35) Yang, Y.; Huang, J.; Yang, X.; Quan, K.; Wang, H.; Ying, L.; Xie, N.; Ou, M.; Wang, K. *J. Am. Chem. Soc.* **2015**, *137* (26), 8340–8343.

(36) Wark, A. W.; Lee, H. J.; Corn, R. M. *Angew Chem Int Ed.* **2008**, *47* (4), 644–652.

(37) Larkey, N. E.; Almlie, C. K.; Tran, V.; Egan, M.; Burrows, S. M. *Anal. Chem.* **2014**, *86* (3), 1853–1863.

(38) Larkey, N. E.; Brucks, C. N.; Lansing, S. S.; Le, S. D.; Smith, N. M.; Tran, V.; Zhang, L.; Burrows, S. M. *Anal. Chim. Acta* **2016**, *909*.

(39) Markham, N. R.; Zuker, M. *Nucleic Acids Res.* **2005**, *33* (SUPPL. 2), 577–581.

(40) Behlke, M. a; Huang, L.; Bogh, L.; Rose, S.; Devor, E. J. *Integr. DNA Technol.* **2005**, 2–4.

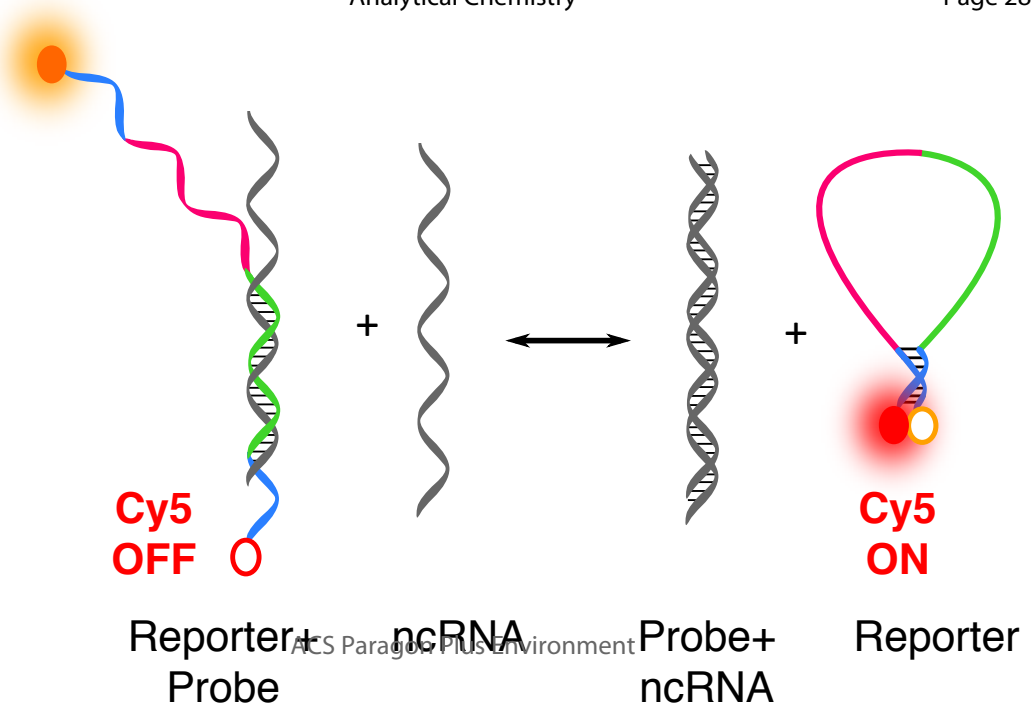
(41) Almlie, C. K.; Larkey, N. E.; Burrows, S. M. *Anal. Methods* **2015**, *7* (17), 7296–7310.

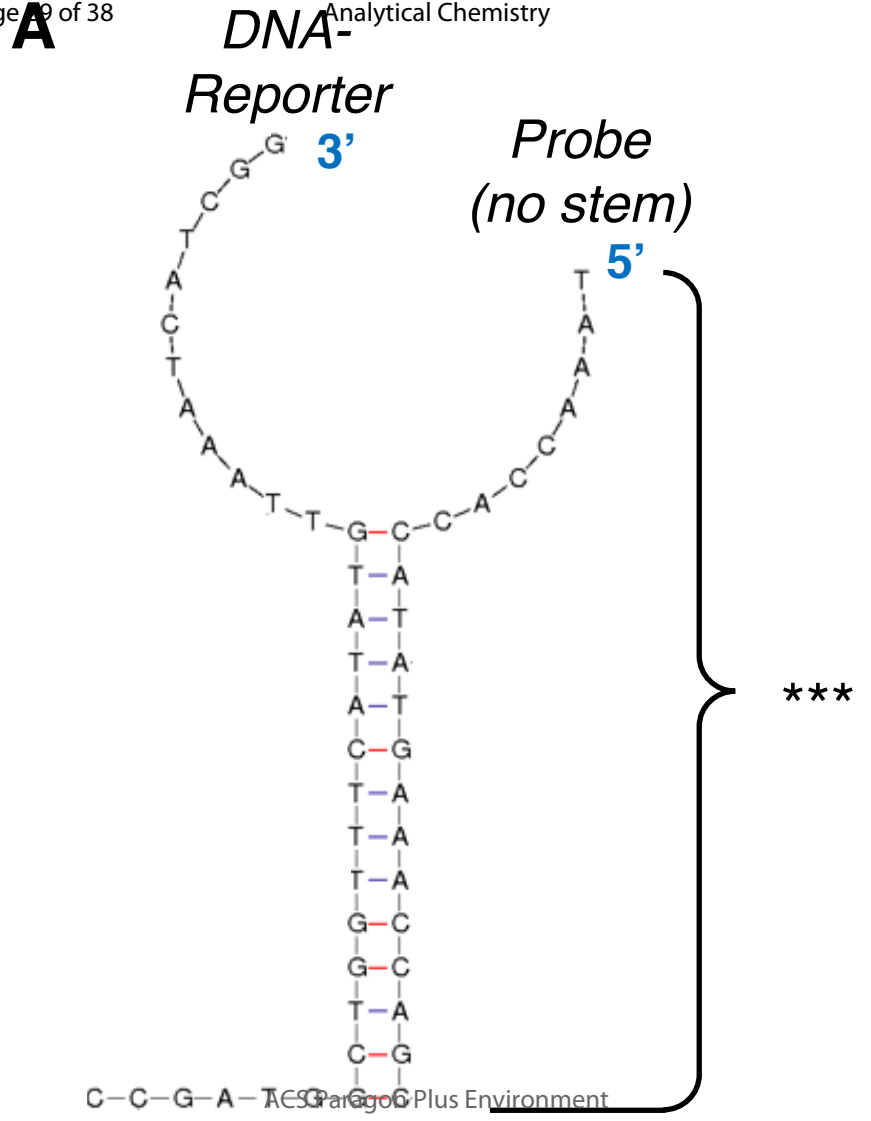
(42) Aken, B. L.; Ayling, S.; Barrell, D.; Clarke, L.; Curwen, V.; Fairley, S.; Fernandez Banet, J.; Billis, K.; García Girón, C.; Hourlier, T.; Howe, K.; Kähäri, A.; Kokocinski, F.; Martin, F. J.; Murphy, D. N.; Nag, R.; Ruffier, M.; Schuster, M.; Tang, Y. A.; Vogel, J.-H.; White, S.; Zadissa, A.; Flicek, P.; Searle, S. M. J. *Database* **2016**, *2016* (September), baw093.

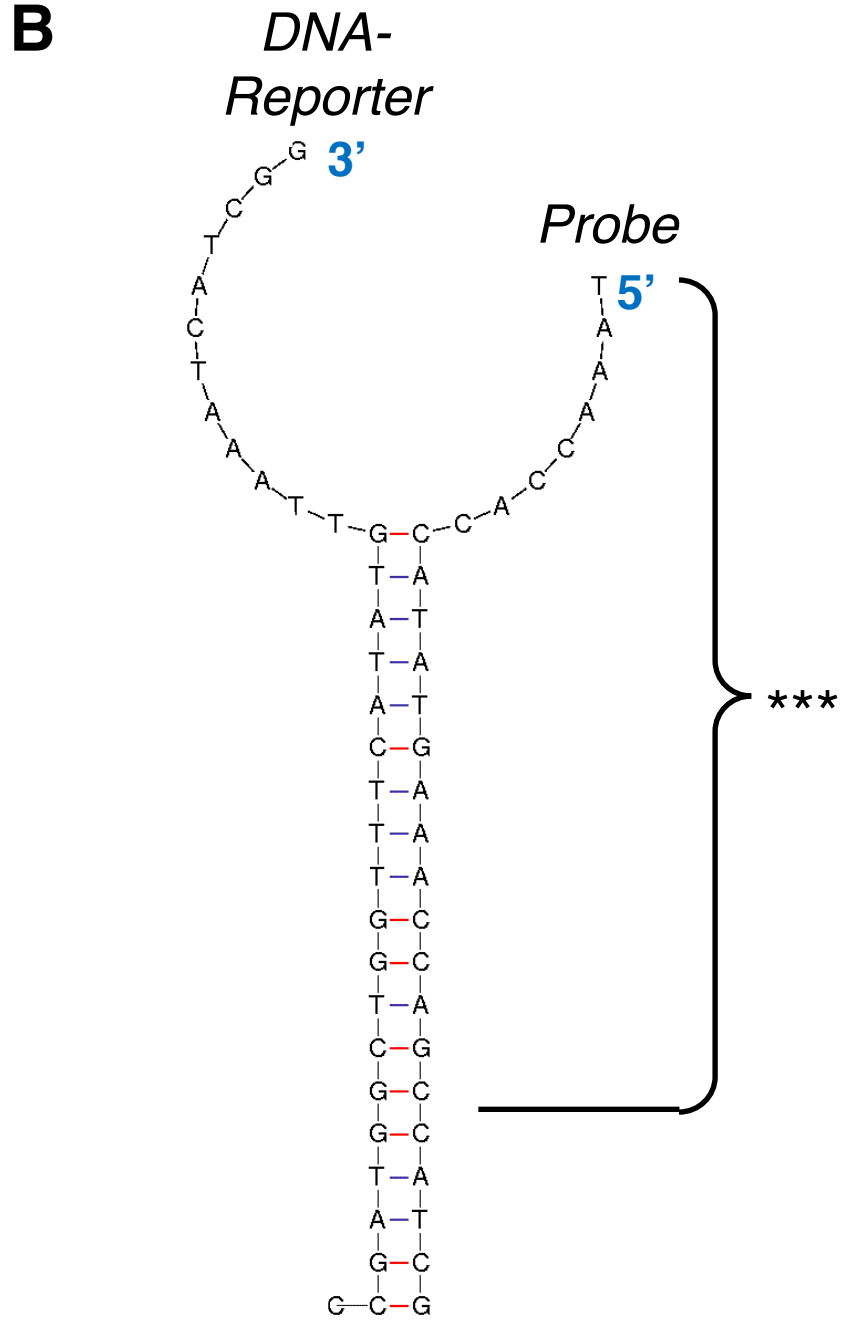
(43) Vogt, V. M. *Methods Enzymol.* **1980**, *65* (C), 248–255.

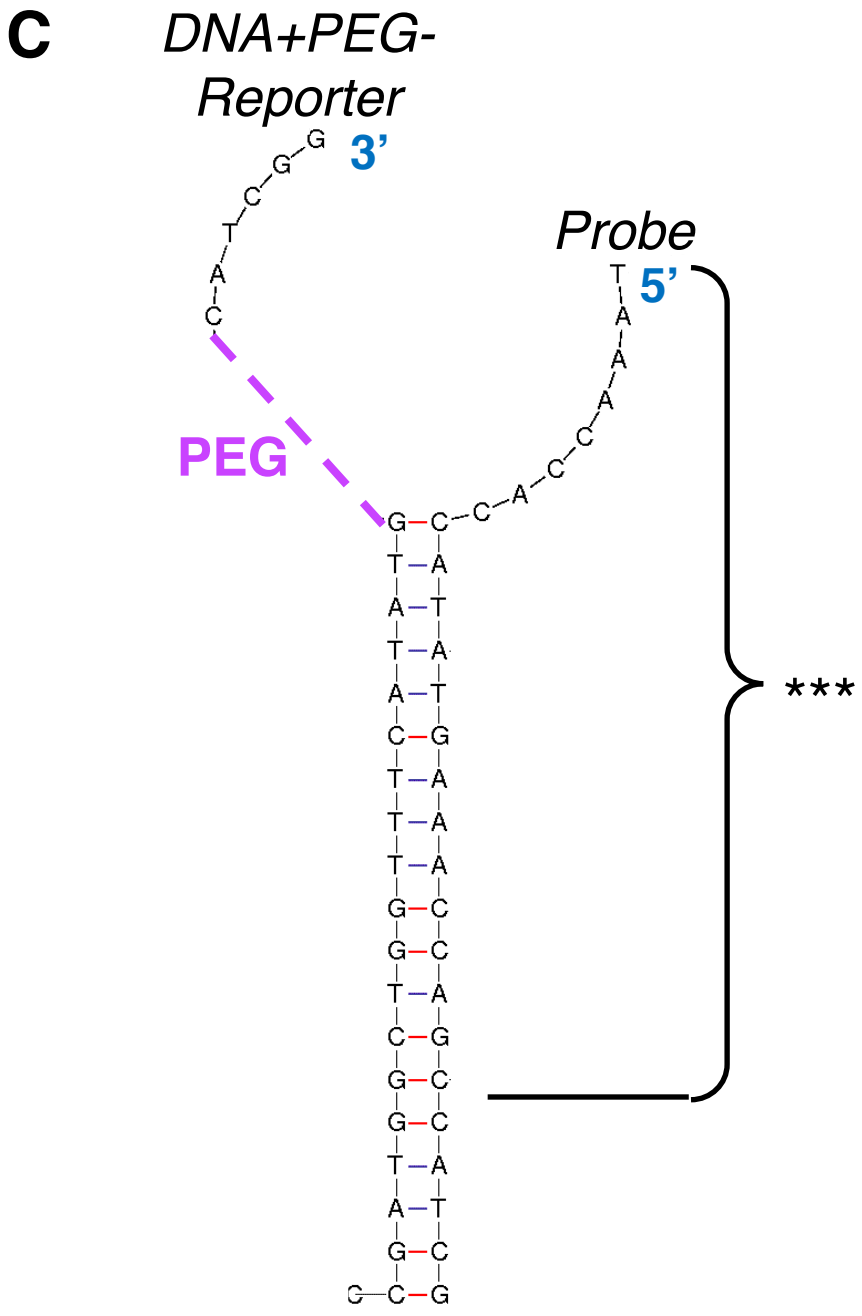
(44) Zhang, L.; Bluhm, A. M.; Chen, K.-J.; Larkey, N. E.; Burrows, S. M. *Nanoscale* **2017**, *9* (4), 1709–1720.

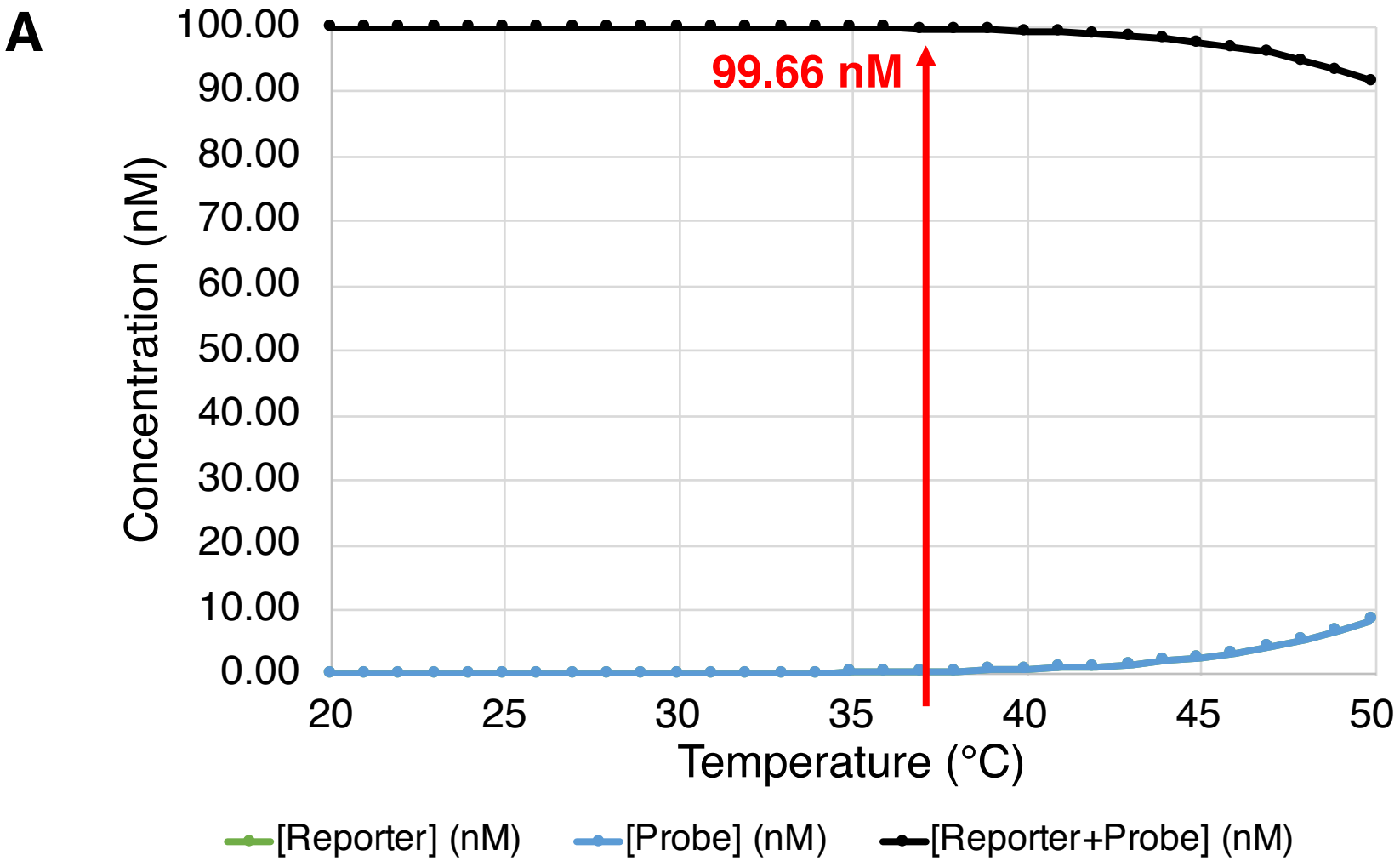
For Table of Contents Only:



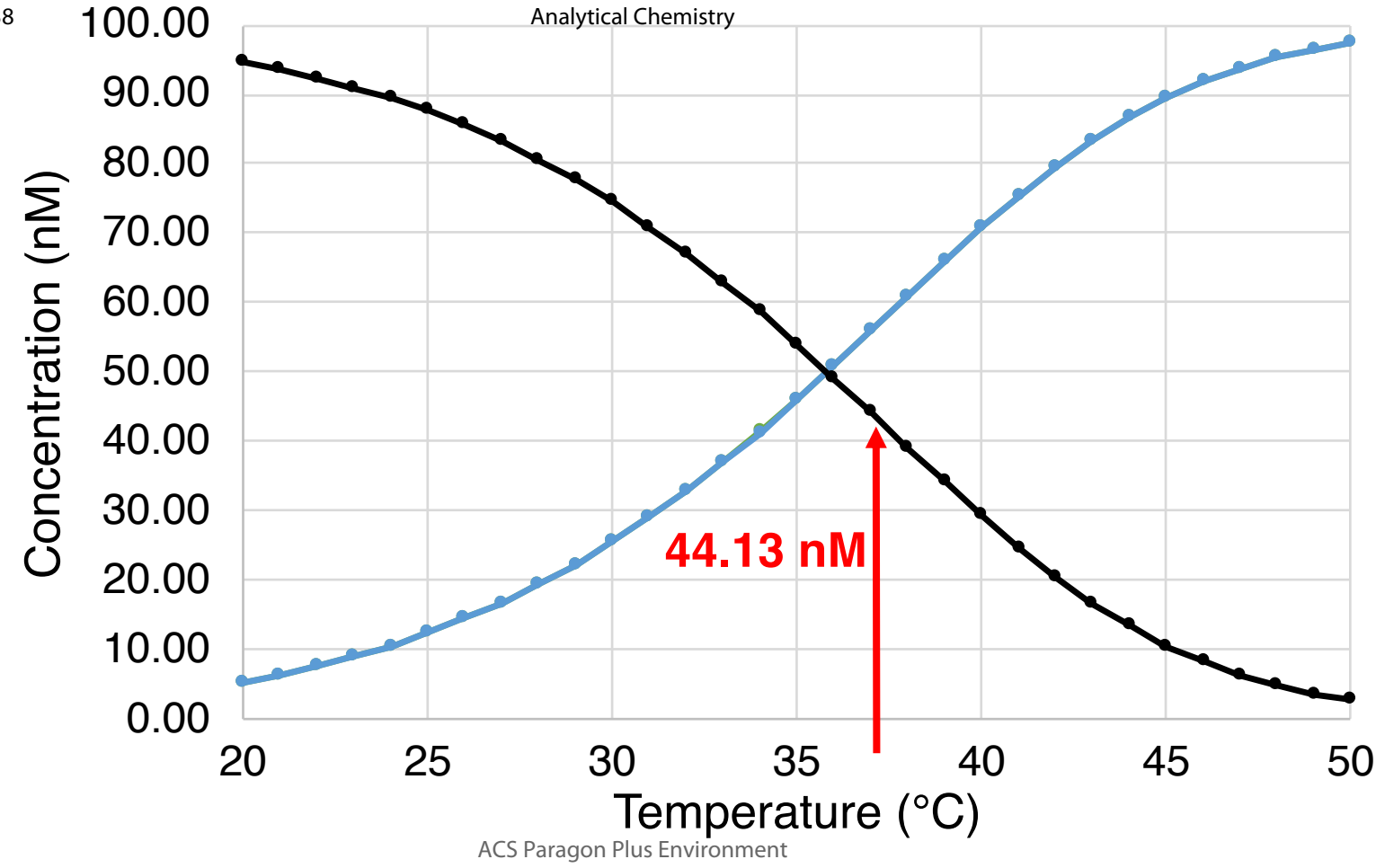


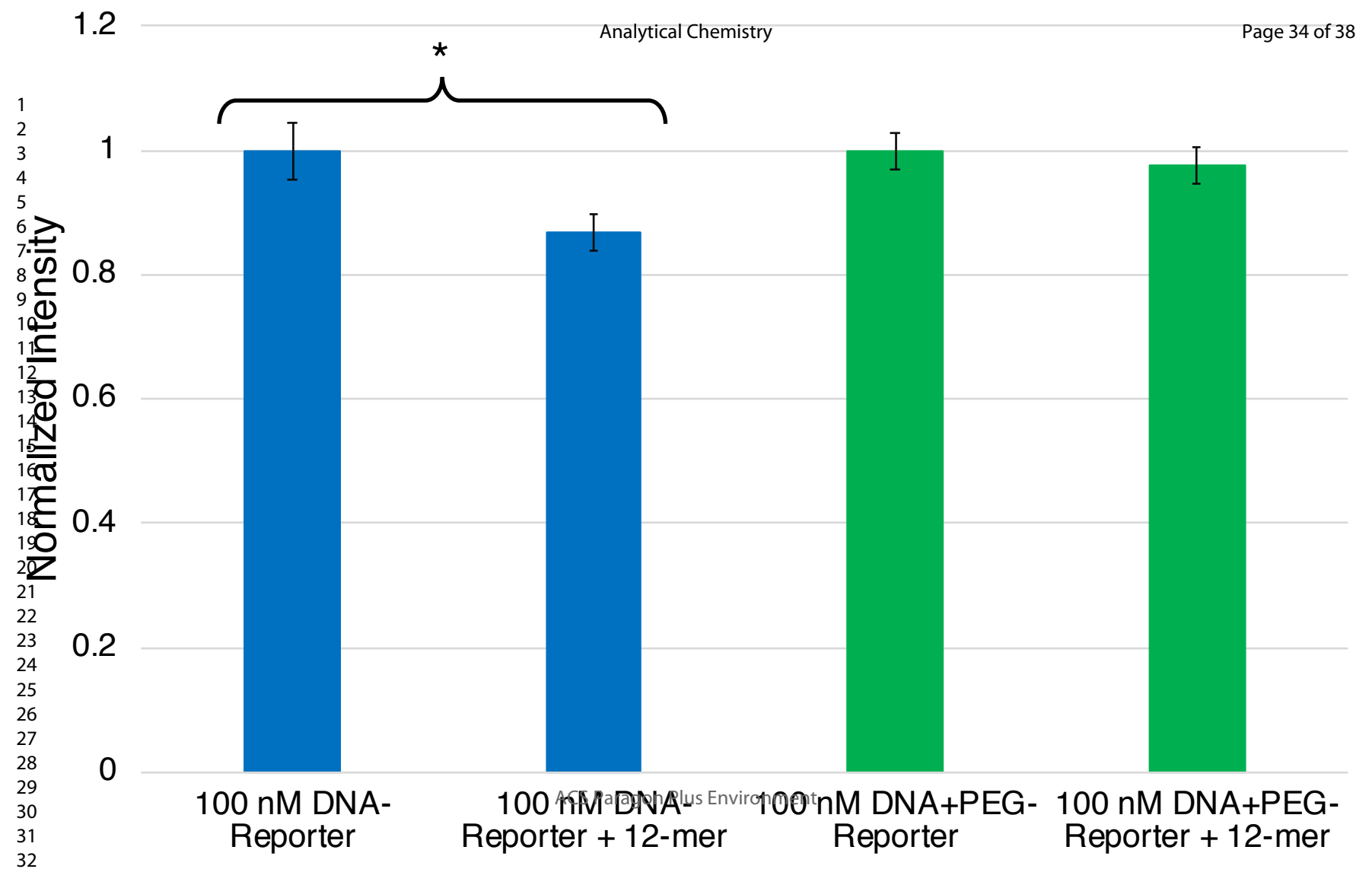

B
DNA-
Reporter
3’
Probe
5’
***

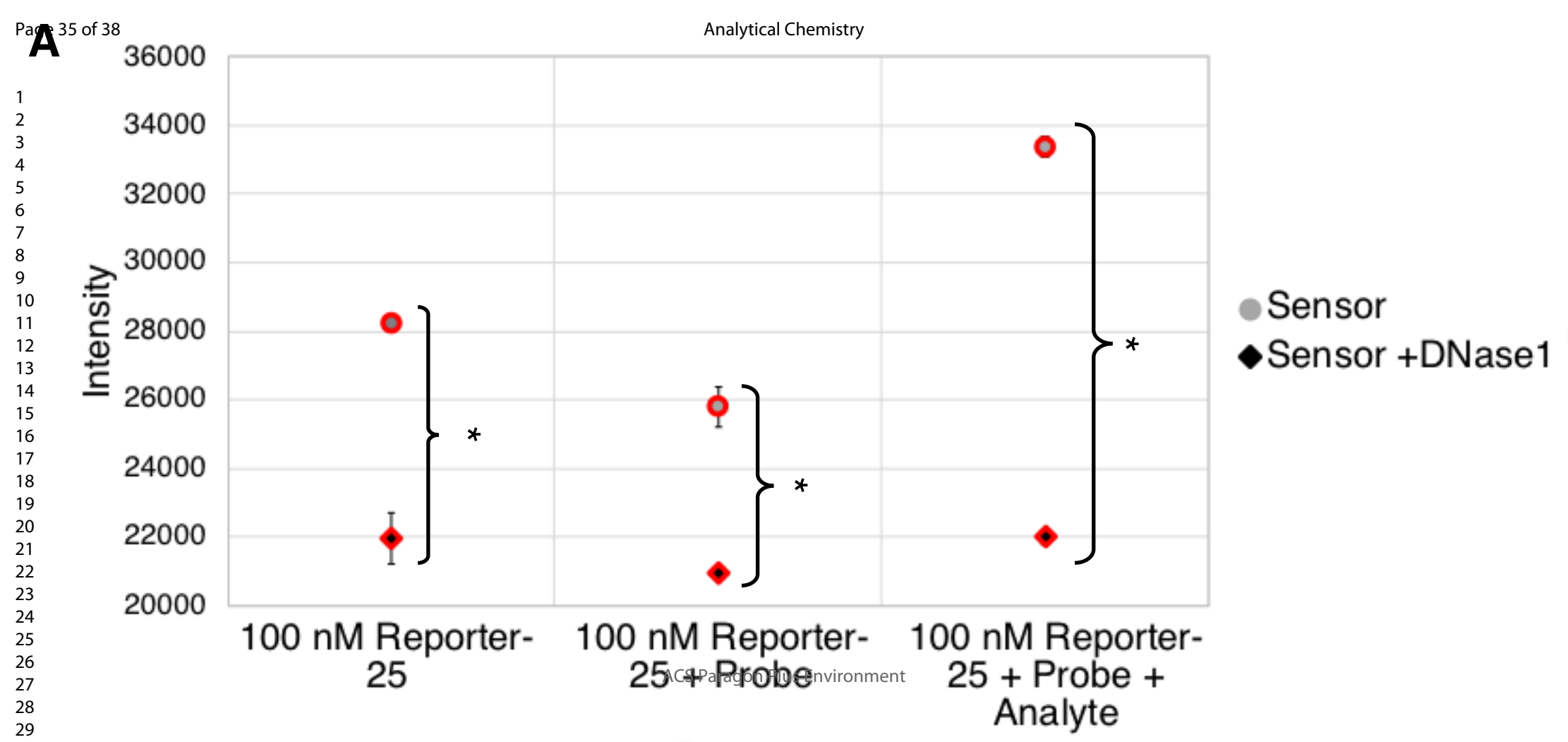

C
DNA+PEG-Reporter
3'
Probe
5'
PEG
***


A
100.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
Concentration (nM)
99.66 nM
20
25
30
35
40
45
50
Temperature (°C)
[Reporter] (nM)
[Probe] (nM)
[Reporter+Probe] (nM)

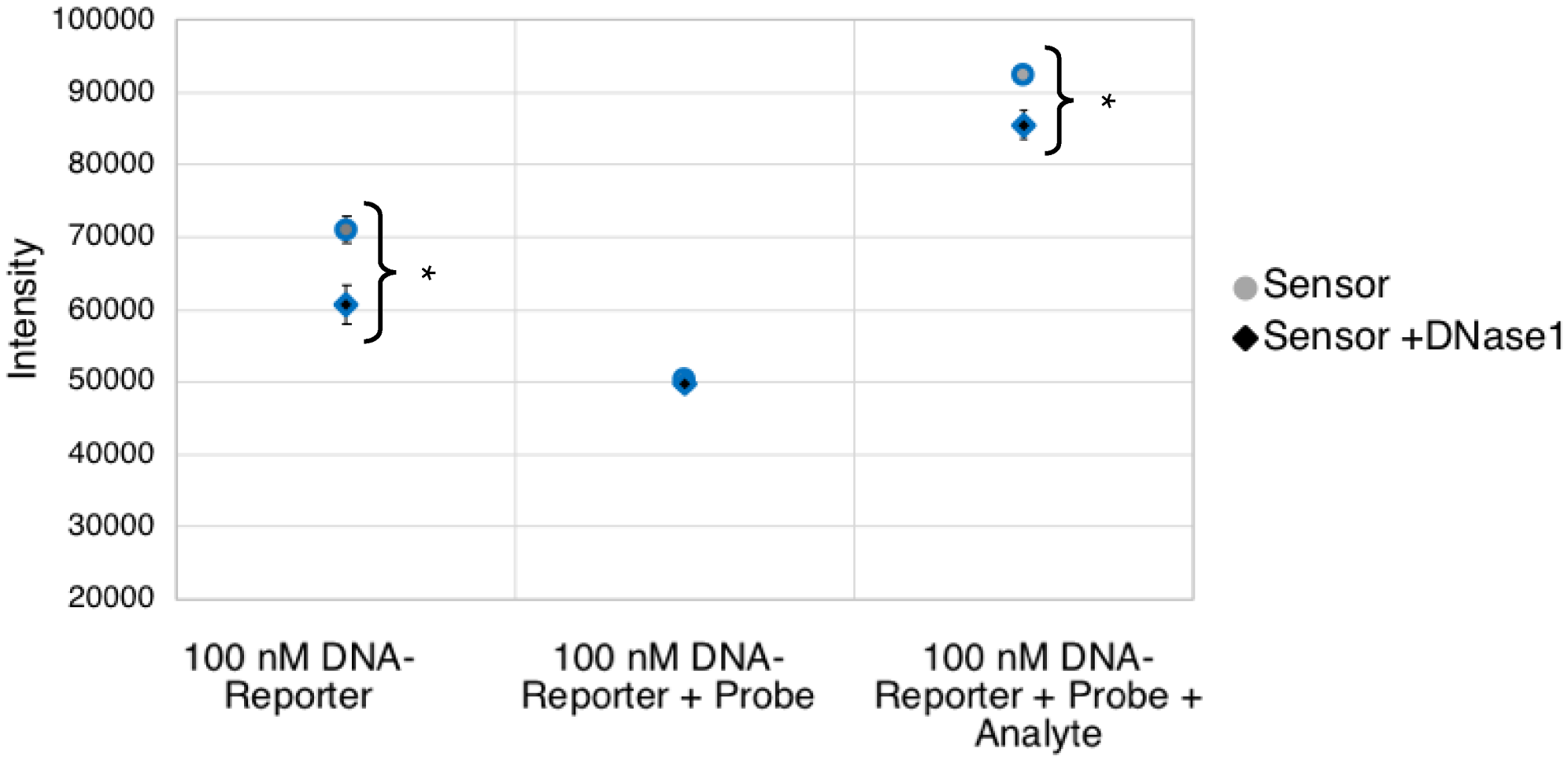
B


100.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
Concentration (nM)
20
25
30
35
40
45
50
Temperature (°C)
44.13 nM
ACS Paragon Plus Environment
[Reporter] (nM)
[Probe (no stem)] (nM)
[Reporter+Probe (no stem)] (nM)

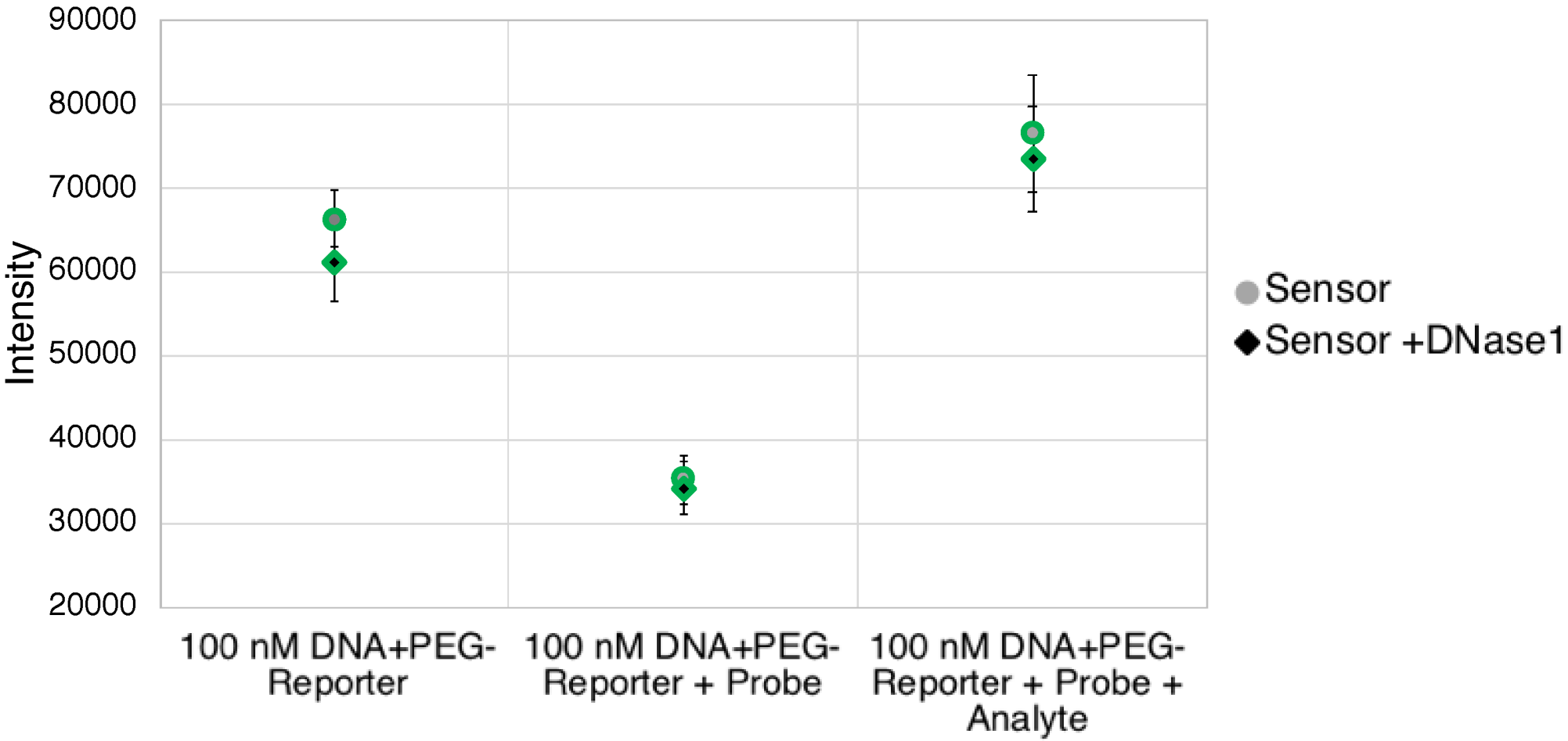




B


100000
90000
80000
70000
60000
50000
40000
30000
20000
Intensity
*
*
100 nM DNA-Reporter
100 nM DNA-Reporter + Probe
100 nM DNA-Reporter + Probe + Analyte
Sensor
Sensor +DNase1

C


90000
80000
70000
60000
50000
40000
30000
20000
Intensity
100 nM DNA+PEG-Reporter
100 nM DNA+PEG-Reporter + Probe
100 nM DNA+PEG-Reporter + Probe + Analyte
Sensor
Sensor +DNase1